

JANEZ ŠTALEC i KOONSTANTIN MOMIROVIĆ  
Fakultet za fizičku kulturu Sveučilišta u Zagrebu

Primljeno 29. 4. 1982.

## JEDNOSTAVAN ALGORITAM ZA ANALIZU HIPOTETSKIH LATENTNIH DIMENZIJA

### SAŽETAK

Predložen je algoritam i napisan program za konfirmativnu faktorsku analizu, koji faktorizira dopustivo singularne matrice korelacija procedurom koja se odvija u dvije faze. U prvoj fazi algoritam formira inicijalnu matricu sklopa jednom modifikacijom multigrupne metode, a u drugoj upotrebljava tu matricu za dobijanje finalne solucije algoritmom koji se temelji na generalnom Guttmanovom modelu faktoriziranja neke matrice kovarijanci.

### 1. UVOD

U primijenjenim kineziolojskim istraživanjima, a tako je u istraživanjima u ostalim antropolojskim znanostima, latentne se dimenzije procjenjuju, u pravilu, na temelju skupova varijabli formiranih u okviru teorijskih modela koji su bili predmet verifikacije u prethodnim, eksplorativno ili konfirmativno orijentiranih analizama latentne strukture manifestnih antropolojskih varijabli. Hipotetska latentna struktura u primijenjenim istraživanjima je stoga eksplicitno definirana, a hipotetske latentne dimenzije pokrivene većim brojem manifestnih varijabli čiji su predmeti mjerenja poznati iz ranijih analiza, ili se s velikom vjerojatnošću mogu pretpostaviti na temelju teoretskih, u pravilu kibernetički formuliranih modela. Zbog toga je procjena latentnih dimenzija u takvim istraživanjima moguća na temelju jednostavnih konfirmativnih algoritama, koji su pogodni ne samo zbog znatne efikasnosti i ekonomičnosti, već i zbog toga što omogućuju vrlo jednostavnu interpretaciju rezultata.

Algoritam KOCHIKI DAOSHI, predložen u ovom radu, i njemu pridruženi program istog imena, pisan u meta jeziku SS (Zakrajšek, Štalc i Momirović, 1974; Momirović, Štalc i Zakrajšek, 1981), faktorizira dopustivo singularne matrice korelacija procedurom koja se odvija u dvije faze. U prvoj fazi formira se, na temelju hipotetske selektorske matrice, inicijalna solucija jednom modifikacijom Holzinger-Thurstonove multigrupne metode (Momirović, 1964; 1966). U drugoj fazi inicijalna matrica sklopa upotrebljava se za formiranje finalne solucije jednim algoritmom koji se temelji na generalnom Guttmanovom modelu faktoriziranja neke matrice kovarijanci (Herak, Gredelj i Momirović, 1982). U finalnoj soluciji algoritam određuje sklop, korelacije i strukturu latentnih dimenzija, formira regresijsku matricu za procjenu tih dimenzija na temelju lijevog pseudoinveza matrice sklopa, određuje komponente varijanci manifestnih i latentnih varijabli, i određuje pouzdanost latentnih dimenzija Cronbachovim postupkom za procjenu generalizabilnosti (Cronbach, Rajaratnam i Gleser, 1963).

### 2. ALGORITAM

Neka je  $E = \{e_i; i=1, \dots, n\}$  uzorak entiteta izvučen nekim slučajnim postupkom iz populacije  $P$ . Neka je

$V = \{v_j; j=1, \dots, m\}$  skup kvantitativnih varijabli, formiran na temelju nekog teoretskog modela o strukturi latentnih dimenzija iz nekog latentnog sustava  $\Omega$ , i neka je  $H = (h_{jp})$ ,  $j=1, \dots, m$ ;  $p=1, \dots, k$  selektorska matrica formirana tako da je  $h_{jp}=1$  ako je manifestna varijabla  $v_j$  mjera neke latentne dimenzije  $w_k$ , i  $h_{jp}=0$  ako varijabla  $v_j$  nije direktna mjera dimenzije  $w_k$ , uz ograničenje da jedna manifestna varijabla može biti direktna mjera jedne, i samo jedne latentne dimenzije  $w_k$  latentnog sustava  $\Omega$ .

Neka je  $Z = (z_{ij})$  matrica standardiziranih podataka dobijena opisom skupa  $E$  nad skupom  $V$ . U matrici

$$R = Z^T Z \quad \text{— biti će interkorelacije varijabli iz } V \text{ na skupu } E$$

Kako je kompozite varijabli iz  $V$  definirane hipotetskom matricom  $H$  moguće definirati operacijom  $K^* = ZH$ . u ma-

$$\text{trici } V^* = Z^T K^* \quad \text{— biti će kovarijance varijabli iz } V$$

i kompozitnih varijabli iz  $K^*$ , a u matrici  $C^* = K^* K^* \quad \text{—}$

$= H^T R H$  kovarijance kompozitnih varijabli iz  $K^*$ . Neka je  $D^2 = \text{diag } C^*$  dijagonalna matrica koja sadrži varijance kompozitnih varijabli. Sada su u matrici  $V = R H D^{-1}$  korelacije varijabli iz  $V$  i latentnih dimenzija definiranih kompozitnim varijablama, u matrici  $C = D^{-1} H^T R H D^{-1}$  interkorelacije latentnih dimenzija, a u matrici  $P = V C^{-1}$  koordinate vektora varijabli iz  $V$  u koordinatnom sustavu definiranom hipotetskim latentnim dimenzijama iz  $\Omega$ .

Pod faktorskim modelom  $Z = \Phi P^T + E$ , gdje je  $\Phi$  matrica faktorskih vrijednosti, a  $E$  matrica rezidualnih vrijednosti, matrica  $\beta = P(P^T P)^{-1}$  je regresijska matrica za procjenu faktorskih vrijednosti za faktore definirane na

$$\text{skupu } \Omega. \text{ ako je } \phi = Z\beta, \text{ to su u matrici } F^* = Z^T \phi \quad \text{— } R\beta$$

kovarijance varijabli iz  $V$  i faktorskih vrijednosti, a u ma-

$$\text{trici } M^* = \phi^T \phi \quad \text{— } \beta^T R \beta \text{ kovarijance faktorskih vrijednosti.}$$

Neka je  $\Delta^2 = \text{diag } M^*$  dijagonalna matrica varijanci ovako procijenjenih faktorskih vrijednosti.

Sada su u matrici  $F=R\beta\Delta^{-1}$  korelacije varijabli iz  $V$  i standardiziranih faktorskih vrijednosti, u matrici  $M=\Delta^{-1}\beta^T R\beta\Delta^{-1}$  korelacije faktorskih vrijednosti, a u matrici  $A=FM^{-1}$  koordinate vektora varijabli u prostoru što ga razapinju vektori faktora. Regresijska matrica za procjenu ovako definiranih dimenzija je  $\Gamma=A(A^T A)^{-1}$ , dakle lijevi pseudo inverz matrice sklopa  $A$ , a konačna procjena vrijednosti entiteta na latentnim dimenzijama je  $\Psi=Z\Gamma$ . Uočimo, uzgred, da algoritam dopušta singularnost matrice  $R$ , ali naravno, pretpostavlja da je hipoteza formulirana u matrici  $H$  takva da ne proizvodi linearno zavisne latentne dimenzije.

$A$  i  $F$  su sigurno faktorske matrice od  $R$ , jer  $AF^T=AMA^T=R\beta(\beta^T R\beta)^{-1}\beta^T R=T$ , što je opći Guttmanov oblik aproksimacije neke matrice kovarijanci; u  $T$  su, stoga, kovarijance varijabli iz  $V$  koje se mogu pripisati latentnim dimenzijama iz  $\Psi$ . U matrici  $N=R-T$  su rezidualne kovarijance varijabli iz  $V$ ; u stvari, matrica  $N$  je

$$\text{procjena matrice } ETE \frac{1}{n}$$

Operacija  $W=AoF$ , gdje je  $o$  simbol Hadamardovog množenja dviju matrica, proizvodi matricu  $W=(w_{jp})$  čiji su elementi, po recima, varijance varijabli iz  $V$  koje se mogu pripisati latentnim dimenzijama, a elementi po kolonama varijance latentnih dimenzija koje se mogu pripisati varijablama iz  $V$ .

Latentne dimenzije iz  $\Psi$  su, naravno, i dalje kompoziti varijabli iz  $V$ . Na temelju općeg izraza za varijancu linearnih kompozita, u matrici  $G=\text{diag}(F^T F)$  su prave varijance latentnih dimenzija. Otuda su Cronbachovi koeficijenti generalizabilnosti tih dimenzija  $\alpha_p=m(m-1)^{-1}(1-g_p^{-1})$ , ako sa  $g_p$  označimo dijagonalne elemente matrice  $G$ .

### 3. PROGRAM

Program KOCHIKI DAOSHI napisan je u verziji 5.2/M programskog sistema SS i pohranjen u programskim bibliotekama FFK\*LIB i SRCE\*SS-MAKRO.

Način na koji se program koristi opisan je u članku Zakrajšeka, Štaleca i Momirovića (1974) i priručniku za korištenje programskog sistema SS (Štalec, Momirović i Zakrajšek, 1981).

KOCHIKI DAOSHI može analizirati skupove od do 250 varijabli na uzorcima od do 10.000 entiteta. Iako nema formalnog ograničenja za broj faktora, osim, naravno, da je taj broj manji od broja varijabli, razumni se rezultati mogu očekivati samo ako je svaka hipotetska latentna dimenzija pokrivena sa najmanje 3 varijable.

KOCHIKI DAOSHI može, osim kvantitativnih varijabli, analizirati i varijable koje su definirane na skupu uređenih kategorija, ako se razumno može pretpostaviti da su te kategorije posljedica djelovanja nekog filtarskog sustava na izlaze iz nekog (nepoznatog) multivarijantno normalnog generatora. U tom slučaju, naredbu STATISTICS u bloku 1 programa treba zamijeniti naredbom NORMALISATION (OLDSC=B, NEWSC=Z, N)

Naravno, ako korisnik programa nije sasvim siguran u ispravnost svoje hipoteze, razborito je da tu hipotezu provjeri i nekim postupkom koji uspoređuje rezultate eksplorativno orijentirane faktorske analize sa rezultatima konfirmativne analize te hipoteze. U tu svrhu, obzirom na algoritme primijenjene u programu KOCHIKI DAOSHI, vjerojatno je najpogodnija metoda koja je implementirana u SS programu ORANGUTAN.

```
OUTPUT (DEVICE=PR)
HEADING (TEXT=KOCHIKI DAOSHI, T)
HEADING (TEXT=HOLZINGER — THURSTONE — GUTTMANOVNA METODA)
HEADING (TEXT=DVOFAZNI ALGORITAM ZA FAKTORIZACIJU MATRICE KORELACIJA, D)
TEXT (TEXT= KOCHIKI DAOSHI)
```

```
*
* *** KOCHIKI DAOSHI**
*
* PROGRAM KOCHIKI DAOSHI FAKTORIZIRA DOPUSTI-
* VO SINGULARNE MATRICE KOORELACIJA PRO-
* CEDUROM KOJA SE ODVIJA U DVIJE FAZE
* U PRVOJ FAZI FORMIRA SE INICIJALNA SOLU-
* CIJA JEDNOM MODIFIKACIJOM HOLZINGER-THURSTO-
* NEOVE MULTIGRUPNE METODE (MOMIROVIĆ, 1964;
* 1966). U DRUGOJ FAZI INICIJALNA MATRICA SKLOPA
* UPOTREBLJAVA SE ZA FORMIRANJE FINALNE SOLU-
* CIJE JEDNIM ALGORITMOM KOJI SE TEMELJI NA GE-
* RALNOM GUTTMANOVOM MODELU FAKTORIZIRANJA
* NEKE MATRICE KOVARIJANCI (HERAK, GREDELJ I
* MOMIROVIĆ, 1982). MODEL, ALGORITAM, PROGRAM
* I PRIMJER PRIMJENE ALGORITMA U ANALIZI REAL-
* NIH PODATAKA OPISANI SU U RADU
*
* STALEC, J. I K. MOMIROVIĆ.
* JEDNOSTAVAN ALGORITAM ZA ANALIZU HIPOTET-
* SKIH LATENTNIH DIMENZIIJA.
* KINEZIOLOGIJA, 12 (1982)
```

```
*
* KORISNIK MORA PRIPREMITI OVE DATOTEKE:
*
* (1) MATRICU U VRIJEDNOSTI ENTITETA NA SKUPU
* KVANTITATIVNIH VARIJABLI
* (2) MATRICU KOJA, U FORMI SELEKTORSKE MATRI-
* CE NAPISANE U TRANSPONIRANOM OBLIKU, DEFINI-
* RA PRIPADANJE VARIJABLI HIPOTETSKIM FAKTORIMA.
*
* BLOK O, ULAZNE OPERACIJE, KONTROLA I SREĐIVA-
* NJE PODATAKA.
```

```
INPUT (SCORE=B)
INPUT (SCORE=TS)
TRANSPOSE (OLD=TS, NEW=S)
```

```
* BLOK 1. DISTRIBUCIJE I PARAMETRI VARIJABLI.
*
HEADING (TEXT=DISTRIBUCIJE I PARAMETRI VARI-
JABLI, D)
STATISTICS (SCORE=B, S, CLASS=9)
```

```

CORRELATION (SCORE=Z)
PRINT (MATRIX=R, TEXT=KORELACIJE VARIJABLI)
DELETE (MATRIX=B)
*
*   BLOK 2. INICIJALNA SOLUCIJA.
*
HEADING (TEXT=INICIJALNA SOLUCIJA, D)
MULT (A=R, B=S, M=SUM)
MULT (A=TS, B=SUM, M=TOT)
SCALE (C=TOT, R=MM)
DIAGMULT (A=SUM, D=TOT, C=0.5 R.M=FF)
INVERSION (R=MM, RINV=INVMM)
MULT (A=FF, B=INVMM, M=P)
DELETE (MATRIX=SUM)
DELETE (MATRIX=TOT)
DELETE (MATRIX=TS)
DELETE (MATRIX=MM)
DELETE (MATRIX=INVMM)
DELETE (MATRIX=FF)
PRINT (MATRIX=P, TEXT=INICIJALNA MATRICA SKLOPA)
*
*   BLOK 3. FINALNA SOLUCIJA.
*
HEADING (TEXT=FINALNA SOLUCIJA, D)
MULT (A=P, TA, B=P, M=PTP)
INVERSION (R=PTP, RINV=TPT)
MULT (A=P, B=TPT, M=BETA)
MULT (A=R, BETA, M=SUM)
MULT (A=BETA, TA, B=SUM, M=TOT)
SCALE (C=TOT, R=M)
DIAGMULT A=SUM, D=TOT, C=0.5, R, M=F)
INVERSION (R=M, RINV=MINV)
MULT (A=F, B=MINV, M=A)
MULT (A=A, B=F, TB, M=RT)
LINEAR (A=R, B=RT, CB=1.0, M=RES)
MULT (A=A, TA, B=A, M=ATA)
INVERSION (R=ATA, RINV=TAT)
MULT (A=A, B=TAT, M=GAMA)
MULT (A=Z, B=GAMA, M=FZ)
PRINT (MATRIX=A, TEXT=SKLOP)
PRINT (MATRIX=M, TEXT=KORELACIJE FAKTORA)
PRINT (MATRIX=F, TEXT=STRUKTURA)
PRINT (MATRIX=GAMA, TEXT=REGRESIJSKI KOEFICI-
JENTI)
STATISTICS (SCORE=FZ, CLASS=9, S, Z=ZZ)
PRINT (MATRIX=RES, TEXT=REZIDUALNE KOVARIJANCE)
HADMULT (A=A, B=F, M=V)
PRINT (MATRIX=V, TEXT=KOMPOONENTE VARIJANCI
VARIJABLI I FAKTORA)
MULT (A=F, B=F, M=VAR)
DIAG (A=VAR, C=-1.0, D=VV)
SCALE (C=VV, R=1)
LINEAR (A=I, B=VV, CB=-1.0, M=ALFA)
PRINT (MATRIX=ALFA, TEXT=POUZDANOST FAKTORA)
*
*   KRAJ PROGRAMA
*
HEADING (TEXT=KRAJ PROGRAMA, D)

```

#### 4. NUMERIČKI PRIMJER

U jednoj analizi utjecaja socijalnih činitelja na razvoj motoričkih sposobnosti (Hošek, Bosnar, Prot i Momirović, osobno saopćenje) formiran je skup od 24 kompozitna motorička testa izabrana tako da pokrivaju hipotetske motoričke sposobnosti različito osjetljive na djelovanje egzogenih činitelja. Ovaj se skup sastojao od 4 testa ritma (RKRBBUB, RKRBNR, RKR3R, RKRPLA), 4 testa sile pokušanih pokreta (DDSELP, DDSFOP, DDSEPK, DDSETR), 4 testa brzine jednostavnih pokreta (BBPDRD, BBPLRD, BBPDRN, BBPDNN), 4 testa repetitivne snage (SRAPT, SRCZTL, SRLDTN, SRAVTR), 4 testa koordinacije (GKLSN, R GAGOSS, GKAVLR, GBKRLP) i 4 testa timinga (TKAZON, TKUPAL, TKLPHV, TKUPRN). Ovi su testovi opisani i njihove metrijske karakteristike i latentna struktura analizirani u radovima Momirovića, Štaleca i Wolfa (1975) i Gredelja, Metikoša, A. Hošek i Momirovića (1975).

Na temelju hipotetske selektorske matrice, formirane u skladu sa intencionalnim predmetima mjerenja ovih testova, dobijena je, na jednom uzorku od 540 odraslih muškaraca, finalna matrica sklopa navedena u tabeli 1. Korelacije latentnih dimenzija dobijenih na kraju algoritma navedene su u tabeli 2.

Kako se vidi iz matrice sklopa<sup>1</sup> algoritam je potvrdio egzistenciju svih 6 hipotetskih faktora. Međutim, test BBPDNN i suviše je slabo saturiran faktorom brzine jednostavnih pokreta, što je u skladu sa rezultatima ranijih analiza ponašanja ovog testa. Test TKAZON, koji je namjerno pogrešno stavljen u grupu testova za procjenu timinga očito ne spada u tu grupu, jer je jednako dobro saturiran i faktorom koordinacije; ovo je, naravno, bilo poznato i iz ranijih analiza latentne strukture koordinacije.

<sup>1</sup> mjere brzine jednostavnih pokreta i mjere opće koordinacije su reflektirane, budući da se radilo o vremenskim testovima

Tabela 1

SKLOP

	KOORIT	DINAMS	BRZJEP	REPSNA	SERPMI	TIMING
RKRBBUB	(.76)	-.04	-.04	.12	.08	.10
RKRBNR	(.73)	-.06	-.09	.11	.21	-.01
RKR3R	(.82)	.11	.03	-.13	-.15	.00
RKRPLH	(.86)	-.03	.10	-.08	-.13	-.07
DDSELP	.10	(.76)	.12	.03	-.17	-.06
DDSFDP	-.06	(.74)	-.21	.02	.11	.06
DDSEPK	-.10	(.77)	.01	-.15	.07	.07
DDSETR	.04	(.62)	.12	.16	-.51	-.05
BBPDRD	-.03	.06	(-.87)	.02	.05	.04
BBPLRD	-.02	.02	(-.87)	.01	.04	.04
BBPDRN	-.04	-.00	(-.84)	-.02	.01	.07
BBPDNN	.10	-.11	-.38	.01	-.19	-.14
SRAPT	.12	.32	.08	(.56)	-.06	-.13
SRCZTL	-.12	-.29	.06	(.85)	-.01	.08
SRLDTN	-.03	.01	-.06	(.79)	.08	-.00

Tabela 1 (nastavak)

SRAVTR	.04	— .00	— .06	(.82)	— .04	.07
GKLSNR	— .03	— .01	.03	.10	(— .81)	— .00
GAGOSS	.26	.06	— .11	— .07	(— .80)	.04
GKAVLR	— .15	.00	— .03	.02	(— .73)	.01
GBKRLP	— .08	— .06	.06	— .03	(— .63)	— .11
TKAZON	.10	.07	— .20	— .01	.34	.34
TKUPAL	— .07	— .06	.11	— .06	— .19	(.87)
TKLPHV	— .01	— .08	.06	.12	.13	(.63)
TKUPRN	.00	.08	.04	— .06	— .23	(.87)

Tabela 2

## KORELACIJE FAKTORA

	KOORIT	DINAMS	BRZJEP	REPSNA	SERPMI	TIMING
KOORIT	1.00	.02	.39	.02	.42	.34
DINAMS	.20	1.00	.21	.55	.21	.16
BRZJEP	.39	.21	1.00	.03	.27	.27
REPSNA	.02	.55	.03	1.00	.20	.13
SERPMI	.42	.21	.27	.20	1.00	.50
TIMING	.34	.16	.27	.13	.50	1.00

Korelacije latentnih dimenzija formiraju konfiguraciju u skladu sa onom koja se mogla očekivati na temelju dosadašnjih analiza relacija primarnih motoričkih sposobnosti; očito se izdvajaju grupe faktora u skladu sa relativnim učešćem sistema za energetska regulaciju, sinergijsku regulaciju i regulaciju formiranja složenih motoričkih struktura.

Prema tome, predloženi algoritam efikasno otkriva realno egzistentne latentne dimenzije, pouzdano procjenjuje saturacije varijabli tim latentnim dimenzijama, i dovoljno je osjetljiv na pogreške u alokaciji varijabli, pa

je stoga vrlo pogodan instrument za konfirmativnu analizu latentne strukture korektno konfiguriranih skupova manifestnih varijabli.

## 5. LITERATURA

1. Cronbach, L. J., M. Rajaratnam and G. C. Gleser: Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology* (1963), 16, 137—163.
2. Gredelj, M., D. Metikoš, A. Hošek i K. Momirović: Model hijerarhijske strukture motoričkih sposobnosti. I. Rezultati dobijeni primjenom jednog neoklasičnog postupka za procjenu latentnih dimenzija. *Kineziologija* (1975), 5, 1—2, 7—81.
3. Herak, M., M. Gredelj i K. Momirović: Algoritam i ponentnim modelom, Zbornik IV Međunarodnog simponentnim modelom. Zbornik IV Međunarodnog simpozija »Kompjuter na Sveučilištu«, Cavtat, 1982.
4. Momirović, K.: Faktorska struktura nekih neurotskih simptoma. Doktorska disertacija, Filozofski fakultet Sveučilišta u Zagrebu, Zagreb, 1964.
5. Momirović, K.: Valjanost psihologijskih mjernih instrumenata. U A. Krković, K. Momirović i B. Petz, »Odbrana poglavlja iz psihometrije i neparametrijske statistike«, DPH i RZZZ, Zagreb, 1966.
6. Momirović, K., J. Štalec i B. Wolf: Pouzdanost nekih kompozitnih testova primarnih motoričkih sposobnosti. *Kineziologija* (1975), 5, 1—2, 169—192.
7. Štalec, J., K. Momirović i E. Zakrajšek: SS-Statistical System. Fakultet za fizičku kulturu Sveučilišta u Zagrebu i Sveučilišni računski centar, Zagreb, 1981.
8. Zakrajšek E., J. Štalec i K. Momirović: SS-programski sistem za multivarijantnu analizu podataka. Zbornik Simpozija »Kompjuter na Sveučilištu«, Sveučilišni računski centar, Zagreb, 1974, str. C8-1 — C8-16.

## A SIMPLE ALGORITHM FOR ANALYSIS OF HYPOTHETICAL LATENT DIMENSIONS

The algorithm was proposed and program was written for a confirmative factor analysis factorizing allowably singular matrices of correlation through a two-phase procedure. In the first phase the algorithm forms the initial pattern matrix through a modification of the multi-group method, while in the second it employs this matrix to obtain the final solutions by an algorithm based on the general Guttman's model of factorization of a matrix of covariances.

Янез Шталец, Константин Момирович

## ПРОСТОЙ АЛГОРИТМ ДЛЯ АНАЛИЗА ГИПОТЕТИЧЕСКИХ ЛАТЕНТНЫХ ИЗМЕРЕНИЙ

Предложен алгоритм и написана программа для конформативного факторного анализа, причем проводится факторизация допустимой сингулярной матрицы корреляций при помощи процедуры, протекающей в двух этапах. На первом этапе алгоритм образует начальную матрицу комплекса при помощи одного варианта мультигруппового метода, а на другом этапе используется полученная матрица для вычисления конечного решения алгоритма, основанного на генеральной модели факторизации некоторой матрицы коварианцы Гуттмана.